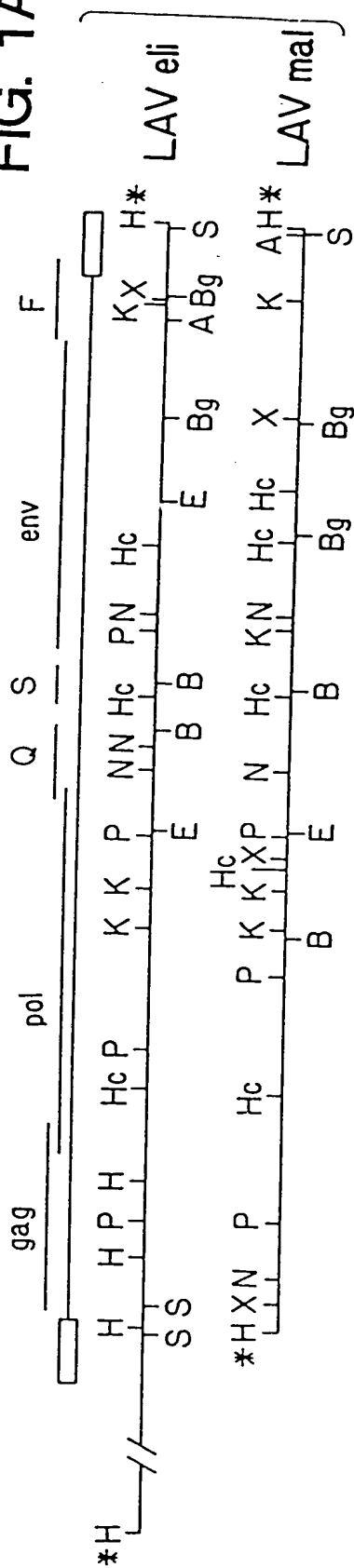


FIG. 1A



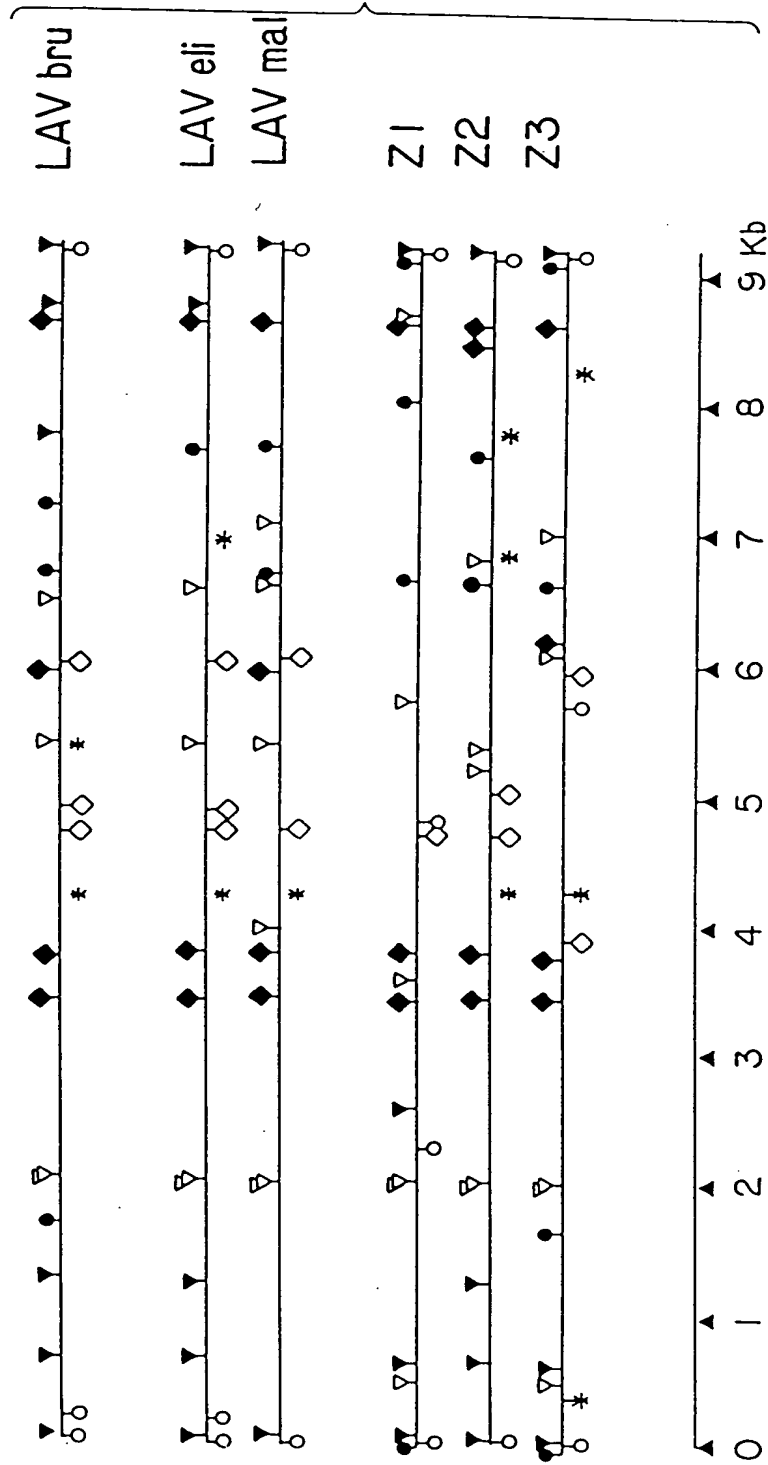
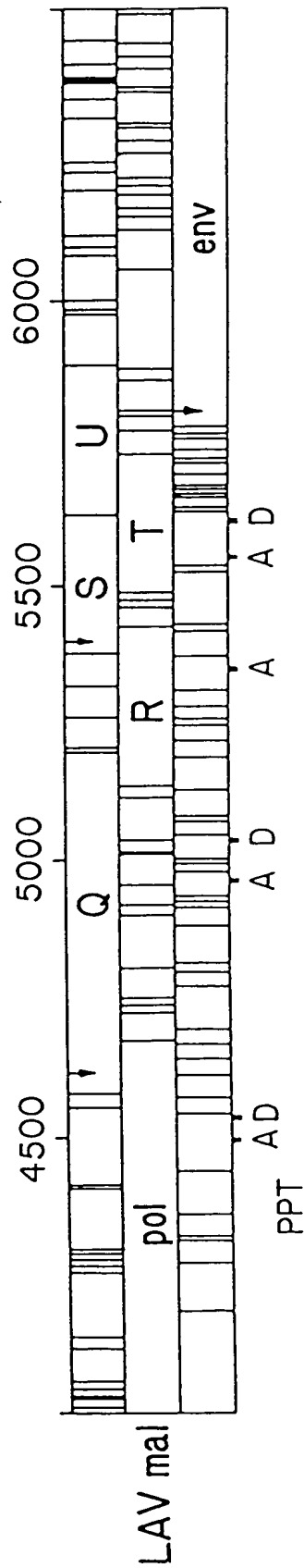
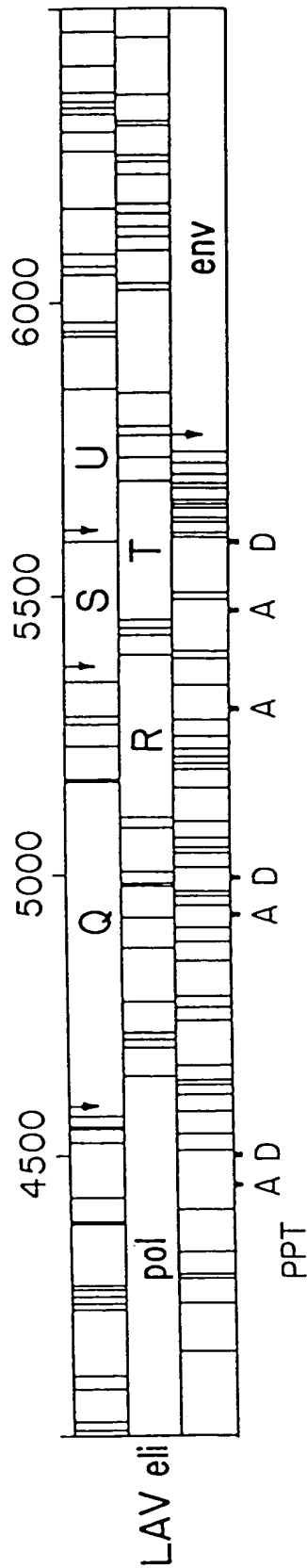
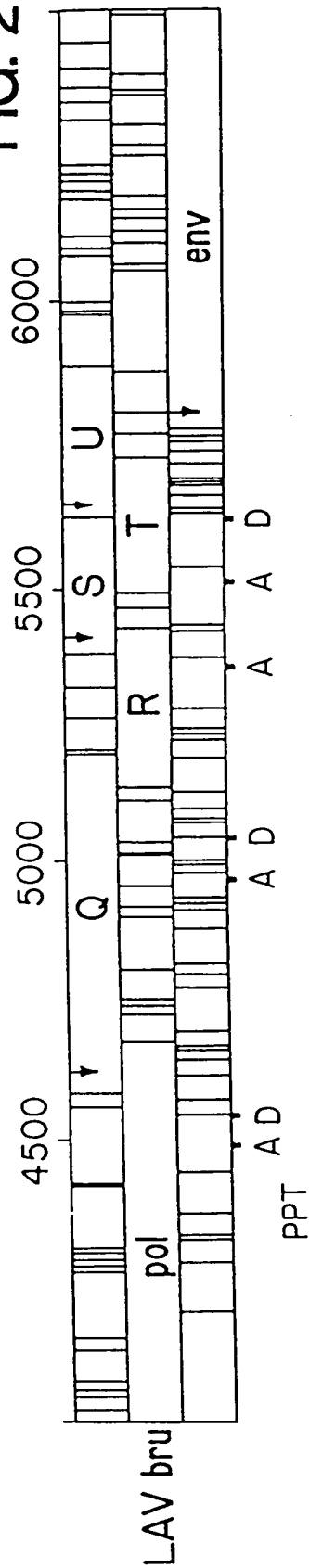


FIG. 1B

FIG. 2



GAG

	10	20	30	40	50	60	70	80
GA6								
LAV BRU	MGARASVLSG	GELDRWEKIR	LRPGGKKKKYK	LKHIVASRE	LERFAVNPGL	LETSEGCQRQI	LGQLQPSLOT	GSEELRSLYN
ARV. 2		K	R	L	L	C	ME	IK
LAV MAL		K	R		Y	Q	I	T
LAV ELI		K	R		L	K	AI	

p25

	90	100	110	120	130	140	150	160
						↓		
LAV BRU	TVATLYCVHQ	RIEIKDTKEA	LDKIEEEQNK	SKKKAAQAAA	-----DTGH	SSQVSONYPI	VQNIQGQMVH	QAISPRTLNA
ARV 2	DV	E			-----AAG	N	L	
LAV MAL	DV	I	RQ	T	AQQA AAA	KN	A	I
LAV ELI	K G DV	E M			-----	N N	L	

	170	180	190	200	210	220	230	240
LAV BRU	WVKVVEEKAF	SPEVIPMFSA	LSCGATPODL	NTMLNTVGGH	QAAMQMLKET	INEEAAEWDK	VHPVHAGPIA	PGQMREPRGS
ARV 2				M	I	D		
LAV MAL	I						L	
LAV ELI	I							

	250	260	270	280	290	300	310	320
LAV BRU	DIAGTTSTLQ	EQIGWMTNPN	PIPVGEIYKR	WIILGLNKIV	RMYSPTSILD	IRQGPKEPFR	DYVDRFYKTL	RAEQASQEVK
ARV 2								D
LAV MAL		S	D		V		F	T
LAV ELI	A	S		V	V			D

FIG. 3A-1

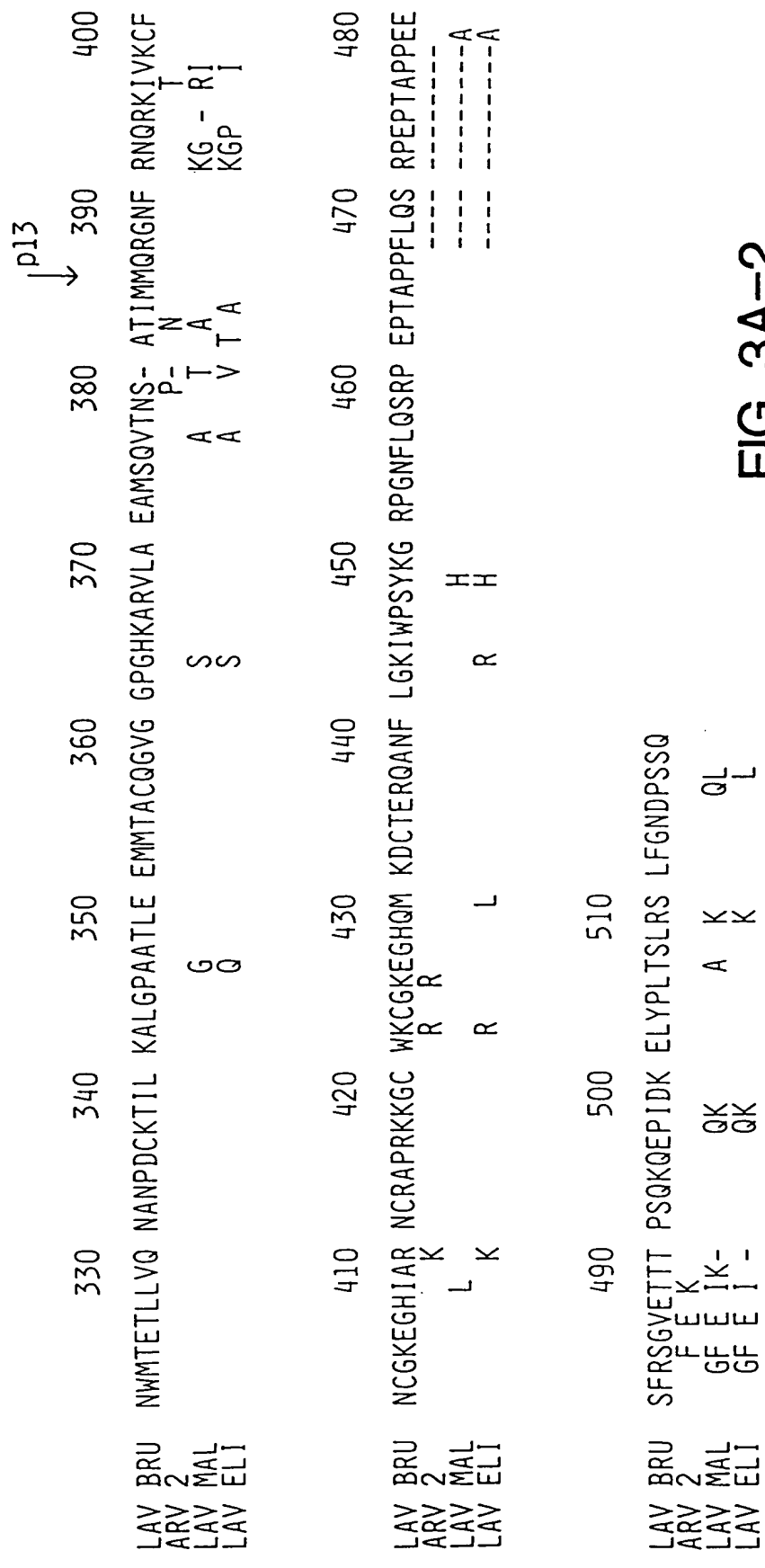


FIG. 3A-2

Central Region: Q

	10	20	30	40	50	60	70	80
LAV BRU	MENRWQVMIV	WQVDRMRIRT	WKS LVKHHMY	VSGKARGWFY	RHHYESPHPR	ISSEVHIPLG	DARLVITTYW	GLHTGERDWH
ARV 2			I K K	K K	T R	V K V	K VR	E K E
LAV MAL		H	K K	KN NR			K	
LAV ELI	K		K			E		

	90	100	110	120	130	140	150	160
LAV BRU	LGQGVSIWR	KKRYSTQVDP	ELADQLIHL	YFDCFSDSA	RKALLGHIVS	PRCEYQAGHN	KVGSLOYLAL	AALITPKKIK
ARV 2	A	K	H	E	KN I YR			T
LAV MAL	H	Q	L	E	Q I	D		T A TR
LAV ELI		R	M	E	I D			T A Q

	170	180	190
LAV BRU	PPLPSVTKLT	EDRWNKPQKT	KGHRGSHTMN GH
ARV 2	K		
LAV MAL	R	Q	
LAV ELI	R	Q R	

FIG. 3B-1

R

	10	20	30	40	50	60	70	80
LAV BRU	MEQAPEDQGP	QREPHNEWTL	ELLEELKNEA	VRHFPRIWLH	GLGQHIYETY	GDTWAGVEAI	IRILQQLLFI	HFRIGCRHSR
ARV 2		Y	R	P	Y			Q
LAV MAL	A		Q	S		E	S	Q
LAV ELI	A	Y A	S	S		V		Q

90

LAV BRU	IGVTQQRARR	-NGASRS
ARV 2	II	R
LAV MAL	I R	- S
LAV ELI	IIR	- S

S (tat)

	10	20	30	40	50	60	70	
LAV BRU	MEPVDPRLEP	WKHPGSQPKT	ACTTCYCKKC	CFHCQVCFTT	KALGISYGRK	KRRQRRRPPQ	GSQTHQVSLS	KQ
ARV 2	N	R	NN	YA	R	A	D	A
LAV MAL	D	N	P NK	Y M	I	N A	DP	P E
LAV ELI	D	N	P NK H	Y P	LN G	G	GA	PIP

FIG. 3B-2

POL

	10	20	30	40	50	60	70	80
LAV BRU	FFREDLAFLO	GKAREFSSEQ	TRANSPTESS	EQTRANSPTR	RELQVWGRDN	NSLSEAGADR	QGTVSFNFPQ	ITLWQRPLVT
ARV 2			---	-----	GE			
LAV MAL	N P	P	---	-----S	R G - KT	T E I	S	V
LAV ELI	N P	G L PK	---	-----S	R - P	KT E		A

	90	100	110	120	130	140	150	160
LAV BRU	IKIGGQLKEA	LLDTGADDTV	LEEMSLPGRW	KPKMIGGIGG	FIKVRQYDQI	LIEICGHKAI	GTVLVGPTPV	NIIGRNLLTQ
ARV 2	R		N K		PV			
LAV MAL	VRV		IN K		P	K I		M
LAV ELI			N K			Q		

	170	180	190	200	210	220	230	240
LAV BRU	IGCTLNFPI	PIETVPVKLK	PGMDGPKVKQ	WPLTEEKIKA	LVEICTEMEK	EGKISKIGPE	NPYNTPVFAI	KKKDKTKWRK
ARV 2								
LAV MAL			R	T	KD	L		
LAV ELI				T	D	R	I	

	250	260	270	280	290	300	310	320
LAV BRU	LVDFRELNKR	TQDFWEVQLG	IPHPAGLKKK	KSVTVLDVGD	AYFSVPLDED	FRKYTAFTIP	SINNETPGIR	YQYNVLPQGW
ARV 2					K			
LAV MAL								
LAV ELI	N							

S

FIG. 3C-1

SECRET 79460

	330	340	350	360	370	380	390	400
LAV BRU	KGSPAIFQSS	MTKILEPFRK	QNPDIVIQY	MDDLYVGSDL	EIGQHRTKIE	ELRQHLLRWG	LTPDKKHQK	EPPFLWMGYE
ARV 2								
LAV MAL		T K	E			E K		
LAV ELI			EM			K E	F R	

	410	420	430	440	450	460	470	480
LAV BRU	LHPDKWTVQP	IVLPEKDSWT	VNDIQKLVGK	LNWASQIYPG	IKVRQLCKLL	RGTKALTEVI	PLTEEALELE	AENREILKEP
ARV 2								
LAV MAL		S	Q D E	A	K K	A	DIV	A
LAV ELI			N ER					

	490	500	510	520	530	540	550	560
LAV BRU	VHGVVYDPSK	DLIAEIQKQG	QGQWTYQIQY	EPFKNLKTGK	YARTRGAHTN	DVKQLTEAVQ	KITTESIVIW	GKTPKFKLP
ARV 2								
LAV MAL	E	V			M	VS	AQ	I
LAV ELI			H	QY	IKS	A	R S	R R

FIG. 3C-2

	890	900	910	920	930	940	950	960
LAV BRU	MNKELKKIIG	QVRDQAEHLK	TAVQMAVFIH	NFKRKGIGG	YSAGERIVDI	IATDIQTKEL	QKQITKIQNF	RVYYRDSRDP
ARV 2	N							KK
LAV MAL	E				I M			N
LAV ELI				RR	I		I	

	970	980	990	1000	1010
LAV BRU	LWKGPAKLLW	KGEGAVVIQD	NSDIKVVPRR	KAKIIRDYGK	Q MAGDDCVAS
ARV 2					RQDED
LAV MAL					G G
LAV ELI		K	V		

FIG. 3D-2

BRU 2 MAL ELI	330	340	350	360	370	380	390	400
	SIRIQRGPR	AFVTIGK-IG	NRQAHCNIS	RAKWNATLKQ	IASKLREQFG	NNKT-IIFKQ	SSGGDPEIVT	HSFNCGGEFF
	Y --	W T RI	DI K	Q N E	VK	- V N	M	R
	G HF--	LY T I-V	DI R Y T N	ETE DK Q	V V	K NS	T	R
	RTP -- L Q	SLY TKS-RS	IIG	Q SK Q	V R	I K P	T	
BRU 2 MAL ELI	410	420	430	440	450	460	470	480
	YCNSTQLFNS	TWFNSTWSTE	CSNNTGSDT	ITLPCRICKQF	INMWQEVGKA	MYAPPISGQI	RCCSNITGLL	LTRDGGNN--
	T N	-----RLN	RTEG K N	I I	I	C	S	T -V
	TSK	Q NGARL-	S STGS	I I	KT	A V	N L	I
	TSG	NI A NNI	TES NSTNTN	Q	K VAGR-	ERN	L	NSSD
								I --
BRU 2 MAL ELI	490	500	510	520	530	540	550	560
	NGGSEIFRPG	GGDMRDNWRS	ELYKYKVVKI	EPLGVAPTKA	KRRVVQREKR	AVGI-GALFL	GFLGAAGSTM	GARSMTLTVQ
	T DT V	I	I R	I	I	V M		V L
	SDN TL		Q		E	I L-		A L
	STN T			R	E	I L-		V

FIG. 3E-2

F

	10	20	30	40	50	60	70	80
LAV BRU	MGGKWSKSSV	VGWPTVRERM	R-----RAEPA	ADGVGAASR-	-----DLEKUG	AITSSNTAAT	NAACAWLEAQ	EE-EEVGFPV
ARV 2	R M G	SAI	RAEP	V	-----			
LAV MAL	I	KI	TP T ET	V QD AVSQ	D C	AA SP N	---- PP	
LAV ELI	I	AI	TM	V	-----	S	D	SD

	90	100	110	120	130	140	150	160
LAV BRU	TPQVPLRRHT	YKAAVDLSHF	LKEKGGLEGL	IHSQRRQDIL	DLWIYUTQGY	FPDWQNYTPC	PGVRYPLTFG	WCYKLVPEP
ARV 2	R	L I		E				
LAV MAL	R	G F	D	VW PK	V		I F	F HS
LAV ELI	R	E L		W KK E	V N I		I	E D

	170	180	190	200	210
LAV BRU	DKVEEANKGE	NTSLLHPVSL	HGMDDPEREV	LEWRFDLSRLA	FHHVARELHP
ARV 2	E	N	E A K	V K K S	M
LAV MAL	EE	NC	E A	K K K N	LR R Q
LAV ELI	QE	TN	E	Q	E K M

FIG. 3F-2

FIG. 4A

A LAVbru vs.		GAG	POL	ENV			
				TOTAL	OMP	TMP	
HTLV-3	USA	512 0/0	1015 0/0	1.3 1.4	507 5/0	1.6 349 0/0	1.1
ARV-2	USA	502 12/2	1003 12/0	3.1 13.0	505 17/10	14.3 350 0/1	11.2
LAVeli	ZAIRE	500 13/1	1002 13/0	5.5 20.7	504 22/14	25.3 349 0/0	13.8
LAVmal	ZAIRE	505 14/7	1002 13/0	7.7 21.7	509 13/10	26.4 350 0/1	14.9
B LAVeli vs.							
LAVmal		505 1/6	1002 0/0	8.4 9.8	509 8/13	23.6 350 0/1	14.3

FIG. 4B

A LAVbru vs.		orf F		central region					
				orf Q		orf R		orf S	
HTLV-3 USA	206 0/0	1.5	192 0/0	0		nd		80 0/0	2.5
ARV-2 USA	210 0/4	12.6	192 0/0	10.0		97 0/1	9.4	81 0/1	15.0
LAVeli ZAIRE	206 1/1	19.4	192 0/0	10.4		96 0/0	11.5	80 0/0	27.5
LAVmal ZAIRE	209 2/5	27.0	192 0/0	12.6		96 0/0	10.4	80 0/0	23.8
B LAVeli vs.									
LAVmal	209 3/6	22.5	192 0/0	12.0		96 0/0	6.3	80 0/0	11.3

GAG

a

120

LAV.BRU	K	A	Q	A	A	A	-	-	-	-	-	D	T
	AAA	GCA	CAG	CAA	GCA	GCT						GAC	ACA
ARV 2	K	A	Q	A	A	A	A	-	-	-	-	G	T
	AAG	GCA	CAG	CAA	GCA	GCT	GCA	GCT				GGC	ACA
LAV.MAL	K	T	Q	A	A	A	A	Q	Q	A	A	A	T
	AAG	ACA	CAG	CAG	GCA	GCA	GCT	GCA	CAG	CAG	GCA	GCT	ACA
LAV.ELI	X	A	Q	A	A	A	A	-	-	-	-	D	T
	AAG	GCA	CAG	CAA	GCA	GCA	GCT	-	-	-	-	GAC	ACA

FIG. 6A-1

	460										470										480										
.BRU																															
N	F	L	Q	S	R	P	E	P	T	A	P	P	F	L	Q	S	R	P	E	P	T	A	P	P							
AAT	TTT	CTT	CAG	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	TTT	CTT	CAG	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	E	GAA	GAG	E	GAA	GAG	
2																															
N	F	L	Q	S	R	P	E	P	T	A	P	P																			
AAT	TTT	CTT	CAG	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	GAA	GAG
.MAL																															
N	F	L	Q	S	R	P	E	P	T	A	P	P																			
AAT	TTC	CTT	CAG	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	GCA	GAG
.ELI																															
N	F	L	Q	S	R	P	E	P	T	A	P	P																			
AAC	TTT	CTC	CAA	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	GCA	GAG

FIG. 6A-2

c		20		30
	LAV.BRU	R M R AGA ATG AGA	- - -	R A E P A CGA GCT GAG CCA GCA
	ARV 2	R M R AGA ATG AGA	R A E P CGA GCT GAG CCA	R A E P A CGA GCT GAG CCA GCA
	LAV.MAL	R I R AGA ATA AGA	- - -	R T P P T CGA ACT CCC CCA ACA
	LAV.ELI	R I R AGA ATA AGA	- - -	R T P P T AGA ACT AAT CCA GCA

FIG. 6A-3

d		40	
	LAV.BRU	V G A A S R GTG GGA GCA GCA TCT CGA	- - - - D - - - - GAC
	ARV 2	V G A V S R GTG GGA GCA GTA TCT CGA	- - - - D - - - - GAC
	LAV.MAL	V G A V S R GTA GGA GCA GTA TCT CAA	D A V S Q GAT GCA GTA TCT CAA D GAT
	LAV.ELI	V G A V S R GTA GGA GCA GTA TCT CGA	- - - - D - - - - GAC

u

LAV. BRU CAG CAC CAC TTG TGG ACA TGG GGC TGG AAA TGG GGC ACC ATG CTC

ARV 2 Q H L W R W G T L L
CAG CAC TTG TGG AGA TGG GGC - - - ACC TTG CTC

LAV. MAL CAA AAC TGG TGG AGA TGG GGC - - - M M L
ATG ATG ATG CTC

LAV.ELI CAA AAC TGG TGG AAA TCG GGC - - - T M L
ATC ATG CTC

150

140
I AV. BRU

L
TTA AAG TGC ACT GAT TTG - GGG AAT GCT ACT
C T D L
K C T D L
L TTA AAG TGC ACT GAT TTG - GGG AAT GCT ACT

M M M E K G E I
ATG ATG ATG GAG - AAA GCA GAG ATA

ARV 2

L N C T D L G K A T N T N S S
 TTA AAT TGC ACT GAT TTG - GGG AAG GCT ACT AAT ACC AAT AGT AGT

W
TGG

K	E	E	I
AAA	GAA	GAA	ATA

K	G	E	I
AAA	GGA	GAA	ATA

FIG. 6B-1

L
TTG AAA ATG GAA ATT - GGA GAA GTG
K M E I V
GGA GAA GTG

G	-	-	-	-	-
GGA	-	-	-	-	-
M					
ATG					

FIG. 6B-2



FIG. 6B-3

h

LAV.BRU

410

F	N	S	T	W
TTT	AAT	AGT	ACT	TGG

420

S	T	E	G	S	N	N	T	E	G
AGT	ACT	GAA	GGG	TCA	AAT	AAC	ACT	GAA	GGA

430

S D T I

AGT GAC ACA ATC

ARV 2

410

F	N	N	T	W
TTT	AAT	AAT	ACA	TGG

420

R	L	N	H	T	K	G
AGG	TTA	AAT	CAC	ACT	AAA	GGA

430

N D T I

AAT GAC ACA ATC

LAV.MAL
C N T S K L F N S T W Q N N G A R L
TGT AAT ACA TCA AAA CTG TTT AAT AGT ACA TGG CAG AAT AAT GGT GCA AGA CTA - - AGT AAT AGC ACA GAG TCA
T G S I
ACT GGT AGT ATC

LAV.ELI
C N T S G L F
TGT AAT ACA TCA GGA CTG TTT
N T N I
AAC ACA AAC ATC

N	S	T	W	N	I	S	A	W	N
AAT	AGT	ACA	TGG	AAT	ATT	AGT	GCA	TGG	AAT

N I T E S N N S T
AAT ATT ACA GAG TCA AAT AAT AGC ACA

FIG. 6B-4

LAV.ELI

→ R
GGTCTCTCTGGTTAGACCAGATTTGAGCCTGGGAGCTCTCTGGCTAGCTAGGGAACCCAC
TGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGT
GTGACTCTGGTAAGTACGATCCCTCAGACCCCTTTAGTCAGAGTGGAAATCTCTAGCA
GTGGCGCCCGAACAGGGACCTGAAAGCGAAAGTAGAACAGAGGAGCTCTCTCGACGCAG
GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGAGCGACTGGTGAGTACGCT
AAAATTTTGGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAA
GlyGlyLysLeuAspLysTrpGluLysIleArgLeuArgProGlyGlyLysLysLysTyr
GCGGGGAAAATTAGATAAATGGGAAAAAATTCGGTTACGGCCAGGAGGAAAGAAAAAAT
ArgLeuLysHisIleValTrpAlaSerArgGluLeuGluArgTyrAlaLeuAsnProGly
ATAGACTAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATATGCACTTAATCCTG
LeuLeuGluThrSerGluGlyCysLysGlnIleIleGlyGlnLeuGlnProAlaIleGln
GCCTTTTAGAAACATCAGAAGGCTGTAAACAAATAATAGGGCAGCTACAACCAGCTATTC
ThrGlyThrGluGluLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHis
AGACAGGAACAGAAGAAGCTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTAC
LysGlyIleAspValLysAspThrLysGluAlaLeuGluLysMetGluGluGluGlnAsn
ATAAAGGAATAGATGTAAAGACACCAAGGAAGCTTTAGAAAAGATGGAGGAAGAGCAA
LysSerLysLysLysAlaGlnGlnAlaAlaAlaAspThrGlyAsnAsnSerGlnValSer
ACAAAAGTAAGAAAAAGGCACAGCAAGCAGCAGCTGACACAGGAACAACAGCCAGGTCA
GlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerPro
GCCAAAATTATCCTATAGTGCAGAACCTACAGGGGCAAATGGTACATCAGGCCATATCAC
ArgThrLeuAsnAlaTrpValLysValIleGluGluLysAlaPheSerProGluValIle
CTAGAAGTTTGAACGCATGGGTAAAAGTAATAGAAGAAAAGGCTTTCAGCCCAGAAGTAA
ProMetPheSerAlaLeuSerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsn
TACCCATGTTTTTCAGCATTATCAGAAGGAGCCACCCCAAGATTTAAACACCATGCTAA
ThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAla
ACACAGTGGGGGGACATCAAGCAGCCATGCAATGCTAAAAGAGACCATCAATGAAGAAG
AlaGluTrpAspArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMetArg
CTGCAGAAATGGGATAGGTTACATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGA
GluProArgGlySerAspIleAlaGlyThrThrSerThrLeuGlnGluGlnIleAlaTrp
GAGAACCAAGGGGAAGTGATATAGCAGGAAGTACTAGTACCCTTCAGGAACAAATAGCAT
MetThrSerAsnProProIleProValGlyGluIleTyrLysArgTrpIleIleValGly
GGATGACAAGTAACCCACCTATCCAGTAGGAGAAATCTATAAAAGATGGATAATTGTGG
LeuAsnLysIleValArgMetTyrSerProValSerIleLeuAspIleArgGlnGlyPro
GATTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTTGGACATAAGACAGGGAC

FIG. 7A

LysGluProPheArgAspTyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAla
CAAAGGAACCTTTTAGAGACTATGTAGACCGTTCTATAAACTCTAAGAGCCGAGCAAG

SerGlnAspValLysAsnTrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAsp
CTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAG

FIG. 7B

LysGluGlyLysIleSerArgIleGlyProGluAsnProTyrAsnThrProIlePheAla
 AAAGGAAGGAAAAATTTCAAGAATTGGGCCTGAAAATCCATACAATACTCCAATATTTGC
 IleLysLysLysAspSerThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLys
 CATAAGAAAAAAGACAGTACCAAGTGGAGAAAATTAGTAGATTTTCAGAGAACTTAATAA
 2300
 ArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAlaGlyLeuLysLys
 GAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCGCATCCTGCAGGGCTGAAAAA
 2400
 LysLysSerValThrValLeuAspValGlyAspAlaTyrPheSerValProLeuAspGlu
 GAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTTCAGTTCCTTAGATGA
 AspPheArgLysTyrThrAlaPheThrIleSerSerIleAsnAsnGluThrProGlyIle
 AGATTTTAGGAAATATACCGCCTTTACCATATCTAGTATAAACAATGAGACACCAGGGAT
 2500
 ArgTyrGlnTyrAsnValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSer
 TAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCGGCAATATTCCAAAG
 SerMetThrLysIleLeuGluProPheArgLysGlnAsnProGluMetValIleTyrGln
 TAGCATGACAAAAATCTTAGAGCCCTTTAGAAAACAAAATCCAGAAATGGTTATCTATCA
 2600
 TyrMetAspAspLeuTyrValGlySerAspLeuGluIleGlyGlnHisArgThrLysIle
 ATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGGACAAAAAT
 2700
 GluLysLeuArgGluHisLeuLeuArgTrpGlyPheThrArgProAspLysLysHisGln
 AGAGAAATTAAGAGAACATCTATTGAGGTGGGGATTTACCAGACCAGATAAAAAACATCA
 LysGluProProPheLeuTrpMetGlyTyrGluLeuHisProAspLysTrpThrValGln
 GAAAGAACCCCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACA
 2800
 SerIleLysLeuProGluLysGluSerTrpThrValAsnAspIleGlnAsnLeuValGlu
 GTCTATAAAACTGCCAGAAAAGGAGAGCTGGACTGTCAATGATATACAGAACTTAGTGGA
 ArgLeuAsnTrpAlaSerGlnIleTyrProGlyIleLysValArgGlnLeuCysLysLeu
 GAGATTAAACTGGGCAAGCCAGATTTATCCAGGAATTAAAGTAAGACAATTATGTAACT
 2900
 LeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeuGlu
 CCTTAGGGGAACCAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAATTAGA
 3000
 LeuAlaGluAsnArgGluIleLeuLysGluProValHisGlyValTyrTyrAspProSer
 ACTGGCAGAAAACAGGGAAATTTTAAAGAACCAGTACATGGAGTGTATTATGACCCATC
 LysAspLeuIleAlaGluIleGlnLysGlnGlyHisGlyGlnTrpThrTyrGlnIleTyr
 AAAAGACTTAATAGCAGAAATACAGAAACAAGGGCACGGCCAATGGACATACCAAATTTA
 3100
 GlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThr
 TCAAGAACCATTATAAAATCTGAAAACAGGAAAGTATGCAAGAATGAGGGGTGCCACAC
 AsnAspValLysGlnLeuAlaGluAlaValGlnArgIleSerThrGluSerIleValIle
 TAATGATGTAAAGCAATTAGCAGAGGCAGTGCAAAGAATATCCACAGAAAGCATAGTGAT
 3200
 TrpGlyArgThrProLysPheArgLeuProIleGlnLysGluThrTrpGluThrTrpTrp
 ATGGGGAAGGACTCCTAAATTTAGACTACCCATACAAAAGGAAACATGGGAAACATGGTG
 3300

FIG. 7C

AlaGluTyrTrpGlnAlaThrTrpIleProGluTrpGluPheValAsnThrProProLeu
GGCAGAGTATTGGCAAGCCACTTGGATTCTGAGTGGGAATTTGTCAATACCCCTCCTTT
ValLysLeuTrpTyrGlnLeuGluLysGluProIleIleGlyAlaGluThrPheTyrVal
AGTAAAATTATGGTACCAGTTAGAGAAGGAACCCATAATAGGAGCAGAACTTTCTATGT
3400
AspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAspArgGly
AGATGGGGCAGCTAATAGAGAGACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGG
ArgGlnLysValValProLeuThrAspThrThrAsnGlnLysThrGluLeuGlnAlaIle
AAGACAGAAAGTTGTCCCTTTGACTGACACGACAAATCAGAAGACTGAGTTACAAGCAAT
3500
AsnLeuAlaLeuGlnAspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla
TAATCTAGCCTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGATTCACAATATGC
3600
LeuGlyIleIleGlnAlaGlnProAspLysSerGluSerGluLeuValAsnGlnIleIle
ATTAGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGAGTTAGTCAATCAAATAAT
GluGlnLeuIleLysLysGluLysValTyrLeuAlaTrpValProAlaHisLysGlyIle
AGAGCAGTTAATAAAAAAGGAAAAGGTTTACCTGGCATGGGTACCAGCACACAAAGGAAT
3700
GlyGlyAsnGluGlnValAspLysLeuValSerGlnGlyIleArgLysValLeuPheLeu
TGGAGGAAATGAACAAGTAGATAAATTAGTCAGTCAAGGAATCAGGAAAGTACTATTTTT
AspGlyIleAspLysAlaGlnGluGluHisGluLysTyrHisAsnAsnTrpArgAlaMet
GGATGGAATAGATAAGGCTCAAGAAGAACATGAGAAATATCACAACAATTGGAGAGCAAT
3800
AlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSerCysAspLys
GGCTAGTGATTTTAACTACCACCCGTGGTAGCAAAAGAAATAGTAGCTAGCTGTGATAA
3900
CysGlnLeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGlyIleTrpGln
ATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCA
LeuAspCysThrHisLeuGluGlyLysValIleLeuValAlaValHisValAlaSerGly
ATTAGATTGTACACACTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGG
4000
TyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeu
CTATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTT
LysLeuAlaGlyArgTrpProValLysValValHisThrAspAsnGlySerAsnPheThr
AAAATTAGCAGGAAGATGGCCAGTAAAGTAGTACATACAGACAATGGCAGCAATTTTCC
4100
SerAlaAlaValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIlePro
CAGTGCTGCAGTTAAGGCCGCCTGTTGGTGGGCAGGTATCAAACAGGAATTTGGAATTCC
4200
TyrAsnProGlnSerGlnGlyValValGluSerMetAsnLysGluLeuLysLysIleIle
CTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTAT
GlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAlaValPheIle
AGGACAGGTAAGAGATCAAGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCAT
4300
HisAsnPheLysArgArgArgGlyIleGlyGlyTyrSerAlaGlyGluArgIleIleAsp
CCACAATTTTAAAAGAAGAAGGGGGATTGGGGGATACAGTGCAGGGGAAAGAATAATAGA

FIG. 7D

IleIleAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleIleLysIleGlnAsn
 CATAATAGCAACAGACATACAACTAAAGAATTACAAAAACAAATTATAAAAATTCAAAA
 4400
 PheArgValTyrTyrArgAspSerArgAspProIleTrpLysGlyProAlaLysLeuLeu
 TTTTCGGGTTTATTACAGAGACAGCAGAGATCCAATTTGGAAAGGACCAGCAAAGCTCCT
 4500
 TrpLysGlyGluGlyAlaValValIleGlnAspLysSerAspIleLysValValProArg
 CTGGAAAGGTGAAGGGGCAGTAGTAATACAAGACAAGAGTGACATAAAGGTAGTACCAAG
 ArgLysValLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAla
 MetGluAsnArgTrpGlnValMetIleValTrpGln
 AAGAAAAGTAAAGATTATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGC
 4600
 SerArgGlnAspGluAsp
 ValAspArgMetArgIleLysThrTrpLysSerLeuValLysHisHisMetTyrValSer
 AAGTAGACAGGATGAGGATTAAACATGGAAAAGTTTAGTAAAACACCATATGTATGTT
 LysLysAlaAsnArgTrpPheTyrArgHisHisTyrGluSerProHisProLysIleSer
 CAAAGAAAGCTAACAGATGGTTTTATAGACATCACTATGAAAGCCCCACCCAAAAATAA
 4700
 SerGluValHisIleProLeuGlyGluAlaArgLeuValIleLysThrTyrTrpGlyLeu
 GTTCAGAAAGTACACATCCCACTAGGAGAAGCTAGACTGGTAATAAAAACATATTGGGGTC
 4800
 HisThrGlyGluArgGluTrpHisLeuGlyGlnGlyValSerIleGluTrpArgLysArg
 TGCATACAGGAGAAAGAGAATGGCATCTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAA
 ArgTyrSerThrGlnValAspProGlyLeuAlaAspGlnLeuIleHisMetTyrTyrPhe
 GGAGATATAGCACACAAGTAGACCCTGGCCTGGCAGACCAACTAATTCATATGTATTATT
 4900
 AspCysPheSerGluSerAlaIleArgLysAlaIleLeuGlyAspIleValSerProArg
 TTGATTGTTTTTCAGAATCTGCTATAAGAAAAGCCATATTAGGAGATATAGTTAGTCCTA
 CysGluTyrGlnAlaGlyHisAsnLysValGlySerLeuGlnTyrLeuAlaLeuThrAla
 GGTGTGAGTATCAAGCAGGACATAACAAGGTAGGATCCCTACAGTATTTGGCACTAACAG
 5000
 LeuIleAlaProLysGlnIleLysProProLeuProSerValArgLysLeuThrGluAsp
 CATTAAATAGCACCAAAACAGATAAAGCCACCTTTGCCTAGTGTTAGGAAGCTAACAGAAG
 5100
 MetGluGlnAlaProAlaAspGlnGlyProGlnArgGluProTyrAsnGluTrpAla
 ArgTrpAsnLysProGlnGlnThrArgGlyHisArgGlySerHisThrMetAsnGlyHis
 ATAGATGGAACAAGCCCCAGCAGACCAGGGGCCACAGAGGGAGCCATACAATGAATGGGC
 Q← LeuGluLeuLeuGluGluLeuLysSerGluAlaValArgHisPheProArgIleTrpLeu
 ATTAGAGCTTTTAGAGGAGCTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATATGGCT
 5200
 HisSerLeuGlyGlnHisIleTyrGluThrTyrGlyAspThrTrpValGlyValGluAla
 CCATAGCTTAGGACAACATATTTATGAACTTATGGGGATACCTGGGTAGGAGTTGAAGC
 IleIleArgIleLeuGlnGlnLeuLeuPheIleHisPheArgIleGlyCysGlnHisSer
 TATAATAAGAATACTGCAACAATTACTGTTTATTCAGAAATTGGGTGTCAACATAG
 5300
 ArgIleGlyIleIleArgGlnArgArgAlaArgAsnGlySerSerArgSer
 MetAspProValAspProAsnLeuGlu
 CAGAATAGGCATTATTCGACAGAGAAGAGCAAGAAATGGATCCAGTAGATCCTAACCTAG
 5400

FIG. 7E

ThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValIleIleArgSerGluAsn
 ACTCAACTGCTGTTGAATGGCAGTCTAGCAGAAGAAGAGGTCATAATTAGATCCGAAAT
 6600
 LeuThrAsnAsnAlaLysAsnIleIleAlaHisLeuAsnGluSerValLysIleThrCys
 CTCACAAACAATGCTAAAAACATAATAGCACATCTTAATGAATCTGTAAAAATTACCTGT
 AlaArgProTyrGlnAsnThrArgGlnArgThrProIleGlyLeuGlyGlnSerLeuTyr
 GCAAGGCCCTATCAAAATACAAGACAAAGAACACCTATAGGACTAGGGCAATCACTCTAT
 6700
 ThrThrArgSerArgSerIleIleGlyGlnAlaHisCysAsnIleSerArgAlaGlnTrp
 ACTACAAGATCAAGATCAATAATAGGACAAGCACATTGTAATATTAGTAGAGCACAATGG
 SerLysThrLeuGlnGlnValAlaArgLysLeuGlyThrLeuLeuAsnLysThrIleIle
 AGTAAACTTTACAACAAGTAGCTAGAAAATTAGGAACCTTCTTAACAAACAATAATA
 6800
 LysPheLysProSerSerGlyGlyAspProGluIleThrThrHisSerPheAsnCysGly
 AAGTTTAAACCATCCTCAGGAGGGGACCCAGAAATTACAACACACAGTTTTAATTGTGGA
 6900
 GlyGluPhePheTyrCysAsnThrSerGlyLeuPheAsnSerThrTrpAsnIleSerAla
 GGGGAATTCTTCTACTGTAATACATCAGGACTGTTTAATAGTACATGGAATATTAGTGCA
 TrpAsnAsnIleThrGluSerAsnAsnSerThrAsnThrAsnIleThrLeuGlnCysArg
 TGAATAATATTACAGAGTCAAATAATAGCACAAACACAAACATCACACTCCAATGCAGA
 7000
 IleLysGlnIleIleLysMetValAlaGlyArgLysAlaIleTyrAlaProProIleGlu
 ATAAACAAATTATAAAGATGGTGGCAGGCAGGAAAGCAATATATGCCCTCCTATCGAA
 ArgAsnIleLeuCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyIle
 AGAAACATTCTATGTTTCATCAAATATTACAGGGCTACTATTGACAAGAGATGGTGGTATA
 7100
 AsnAsnSerThrAsnGluThrPheArgProGlyGlyGlyAspMetArgAspAsnTrpArg
 AATAATAGTACTAACGAGACCTTTAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA
 7200
 SerGluLeuTyrLysTyrLysValValGlnIleGluProLeuGlyValAlaProThrArg
 AGTGAATTATATAAATATAAGGTAGTACAAATTGAACCACTAGGAGTAGCACCCACCAGG
 AlaLysArgArgValValGluArgGluLysArgAlaIleGlyLeuGlyAlaMetPheLeu
 GCAAAGAGAAGAGTGGTGGAAAGAGAAAAAGAGCAATAGGATTAGGAGCTATGTTCTT
 7300
 GlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaArgSerValThrLeuThrValGln
 GGGTTCTTGGGAGCAGCAGGAAGCACGATGGGCGCACGGTCAGTGACGCTGACGGTACAG
 AlaArgGlnLeuMetSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGlu
 GCCAGACAATTAATGTCTGGTATAGTGCAACAGCAAAACAATTTGCTGAGGGCTATAGAG
 7400
 AlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgIle
 GCGCAACAGCATCTGTTGCAACTCACGGTCTGGGGCATTAAACAGCTCCAGGCAAGAATC
 7500
 LeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
 CTGGCTGTGGAAGATACCTAAAGGATCAACAGCTCCTAGGAATTTGGGGTTGCTCTGGA

FIG. 7G

LysHisIleCysThrThrAsnValProTrpAsnSerSerTrpSerAsnArgSerLeuAsn
 AACACATTTGCACCACTAATGTGCCCTGGAAGCTCTAGTTGGAGTAATAGATCTCTAAAT
 7600
 GluIleTrpGlnAsnMetThrTrpMetGluTrpGluArgGluIleAspAsnTyrThrGly
 GAGATTTGGCAGAACATGACCTGGATGGAGTGGGAAAGAGAAATTGACAATTACACAGGC
 LeuIleTyrSerLeuIleGluGluSerGlnThrGlnGlnGluLysAsnGluLysGluLeu
 TTAATATATAGCTTAATTGAGGAATCGCAGACCCAGCAAGAAAAGAATGAAAAAGAATTG
 7700
 LeuGluLeuAspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrGlnTrpLeuTrp
 TTGGAATTGGACAAGTGGGCAAGTTTGTGGAATTGGTTTAGCATAACACAATGGCTGTGG
 7800
 TyrIleLysIlePheIleMetIleIleGlyGlyLeuIleGlyLeuArgIleValPheAla
 TATATAAAATATTTCATAATGATAATAGGAGGCTTGATAGGTTTAAGAATAGTTTTTGTCT
 ValLeuSerLeuValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrLeu
 GTGCTTTCTTTAGTAAATAGAGTTAGGCAGGGGATACTCACCTCTGTCTGTTTCAGACCCCTC
 7900
 LeuProAlaProArgGlyProAspArgProGluGlyThrGluGluGluGlyGlyGluArg
 CTCCAGCCCCGAGGGGACCCGACAGGCCCGAAGGAACAGAAGAAGAAGGTGGAGAGCGA
 GlyArgAspArgSerValArgLeuLeuAsnGlyPheSerAlaLeuIleTrpAspAspLeu
 GGCAGAGACAGATCCGTGAGATTGCTGAACGGATTCTCGGCACCTTATCTGGGACGACCTG
 8000
 ArgSerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuIleLeuIleAlaValArg
 CGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTAATCTTAATTGCAGTGAGG
 8100
 IleValGluLeuLeuGlyArgArgGlyTrpAspIleLeuLysTyrLeuTrpAsnLeuLeu
 ATTGTAGAACTTCTGGGACGCAGGGGGTGGGACATCCTCAAATATCTGTGGAATCTCCTA
 GlnTyrTrpSerGlnGluLeuArgAsnSerAlaSerSerLeuPheAspAlaIleAlaIle
 CAGTATTGGAGTCAGGAAGTGGGAAACAGTGCTAGTAGCTTGTGTTGATGCCATAGCAATA
 8200
 AlaValAlaGluGlyThrAspArgValIleGluIleIleGlnArgAlaCysArgAlaVal
 GCAGTAGCTGAGGGGACAGATAGAGTTATAGAAATAATACAAAGAGCTTGAGAGCTGTT
 LeuAsnIleProArgArgIleArgGlnGlyLeuGluArgSerLeuLeu
 CTTAACATACCCAGAAGAATAAGACAGGGCTTAGAAAGGTCTTTACTTTAAATGGGTGG
 8300
 LysTrpSerLysSerSerIleValGlyTrpProAlaIleArgGluArgIleArgArgThr
 CAAATGGTCAAAAAGTAGTATAGTGGGATGGCCTGCTATAAGGGAAAGAATAAGAAGAAC
 8400
 AsnProAlaAlaAspGlyValGlyAlaValSerArgAspLeuGluLysHisGlyAlaIle
 TAATCCAGCAGCAGATGGGGTAGGAGCAGTATCTCGAGACCTGGAAAAACATGGGGCAAT
 ThrSerSerAsnThrAlaSerThrAsnAlaAspCysAlaTrpLeuGluAlaGlnGluGlu
 CACAAGTAGCAATACAGCAAGTACTAATGCTGACTGTGCCTGGCTAGAAAGCACAGAAGA
 8500
 SerAspGluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLys
 GAGCGACGAGGTGGGCTTTCCAGTCAGACCCCAGGTACCTTTAAGACCAATGACTTACAA
 GluAlaLeuAspLeuSerHisPheLeuLysGluLysGlyGlyLeuGluGlyLeuIleTrp
 AGAAGCTCTAGATCTCAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTTG
 8600

FIG. 7H

09767138.012304

SerLysLysArgGlnGluIleLeuAspLeuTrpValTyrAsnThrGlnGlyIlePhePro
GTCCAAAAAGAGACAAGAGATCCTTGATCTTTGGGTCTACAACACACAAGGCATCTTCCC
8700
AspTrpGlnAsnTyrThrProGlyProGlyIleArgTyrProLeuThrPheGlyTrpCys
TGATTGGCAAACCTACACACCAGGGCCAGGGATCAGATATCCACTAACCTTTGGATGGTG
TyrGluLeuValProValAspProGlnGluValGluGluAspThrGluGlyGluThrAsn
CTACGAGCTAGTACCAGTTGATCCACAGGAGGTAGAAGAAGACACTGAAGGAGAGACCAA
8800
SerLeuLeuHisProIleCysGlnHisGlyMetGluAspProGluArgGlnValLeuLys
CAGCTTGTTACACCCTATATGCCAGCATGGAATGGAGGACCCGGAGAGACAAGTGTTAAA
TrpArgPheAsnSerArgLeuAlaPheGluHisLysAlaArgGluMetHisProGluPhe
ATGGAGATTTAACAGCAGACTAGCATTGAGCACAAGGCCCGAGAGATGCATCCGGAGTT
8900
TyrLysAsn
CTACAAAACTGATGACACCGAGCTTTCTACAAGGGACTTTCGCTGGGGACTTTCCAGG
9000
GAGGCGTGGACTGGGCGGGACTGGGGAGTGCTAACCCTCAGATGCTGCATATAAGCAGC
TGCTTTTGGCTGTACTGGTCTCTCTGGTTAGACCAGATTTGAGCCTGGGAGCTCTCTG
U3 ← → R
9100
GCTAGCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCTTGAGTGCTTCAA
B ←

FIG. 71